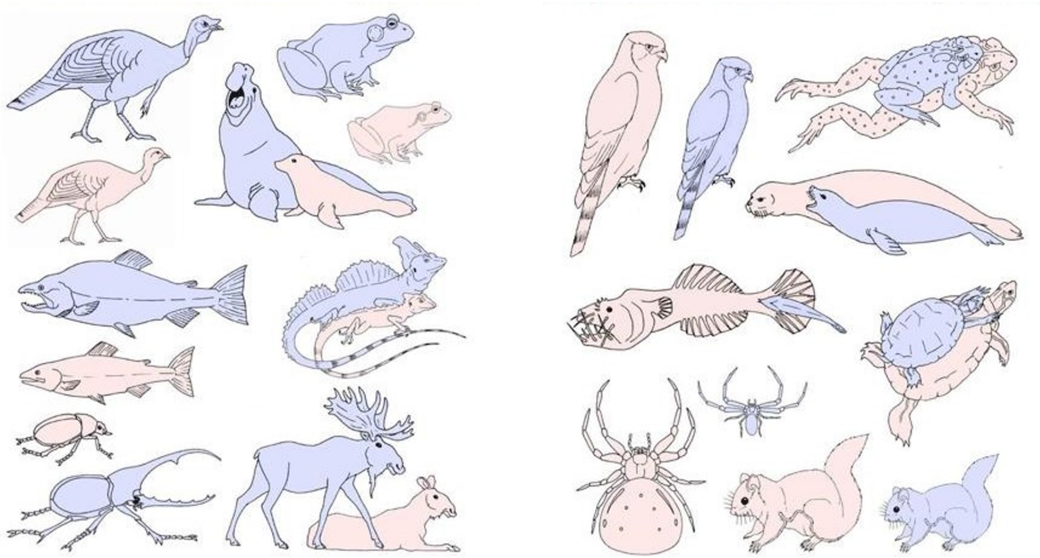
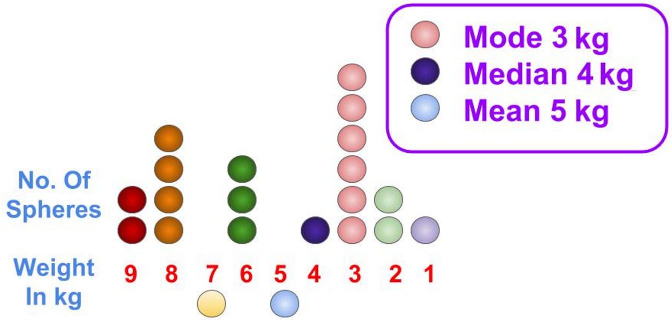
What does "ON AN AVERAGE" mean?

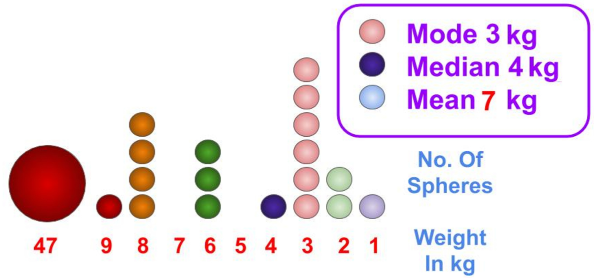


In many animals, the male is much larger and heavier than the female. {Picture above} Examples other than the humans are the deer and the lion. In many other animals, this is reversed and the female is larger and heavier. But it is not correct to say that all male deer are larger than all female deer. All baby animals are small. But even among fully grown adults, there can be some females which are larger than some males. So we come to the conclusion that male deer are larger, "on an average". When we say some female deer are larger, an obvious question arises? How many? How much larger? These are questions that are common for all experimental observations. If we repeat an experiment the result will not be exactly the same. So to compare two measurements we have to also compare these changes. So we need a numerical value for the average. Everyone knows that the

average of two numbers is obtained by adding them and dividing by 2. The average of 2,4 is 3. We also know that the average of 2,4,6 is 2+4+6=12/3=4. But the average of not only 2,4 but 1.5 also is 3. So the average value cannot tell us what the original numbers are.

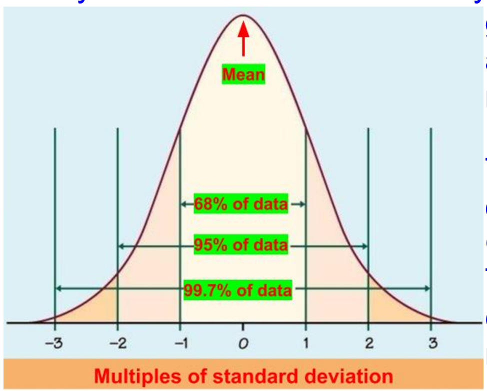
In the picture, {Picture below left} we have 19 spheres varying in weight from 1 to 9 kg. The average is 5kg. But there is no sphere with a weight of 5 kg. Six of the nineteen spheres weigh 3 kg each, That is the most common value and this is called the "mode". The mode gives better information about the actual spheres than the average. If we arrange the balls in an ascending order of weight, nine spheres are lighter than a sphere of 4 kg and nine spheres are heavier. Once again, such a number called the "median", gives a bit more information about the actual spheres than the average. But changing even one sphere changes the average. Median and mode may not change. For example, if one sphere of 9 kg weight is replaced by another weighing 47 kg the median and mode are not changed. {Picture below right} So we cannot know that one sphere much heavier than the rest has been added. But the average changes to 7 kg. So the average, calculated from numbers is more useful than the median and mode determined by observing the collection which is





also called the distribution. There is a difference between the weights of individual spheres and the average. A useful number called the standard deviation is calculated from these differences. This is obtained by first squaring the deviations, finding the average of the deviation squares, dividing by the number of deviations and then finding the square root. Squaring ensures that the data with large increases standard deviation by a large amount. The standard deviation for the original distribution of spheres is 2.8. Adding the extremely heavy sphere increases this to 8. The very large value of standard deviation as compared to the average shows that the distribution cannot be trusted.

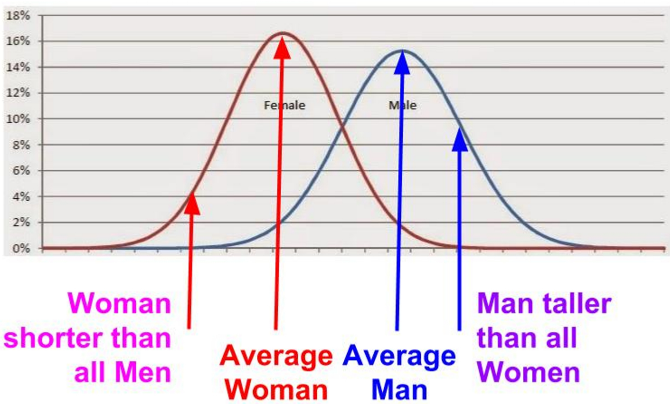
It is completely unscientific to find an excuse and reject items with the large deviation. For example we should not simply remove the 47 kg sphere. We have to identify how that came into the distribution, and repeat the experiment taking precautions to ensure that the deviations are small. That is the scientific approach. Those who challenge modern science invariably find arguments to justify rejecting observations which do not agree with their theory. As mentioned earlier only modern medicine relies on the comparison of identical groups. This ensures that individual measurements with a lot of deviation are not selectively rejected after the results are obtained.



The standard deviation is most useful whenever the data distribution is a Gaussian. Carl Gauss, the very famous German mathematician and physicist first described it. The distribution has the shape of a bell and it is often called the bell curve. {Picture left} The mean, median and mode of a Gaussian distribution are always equal. More

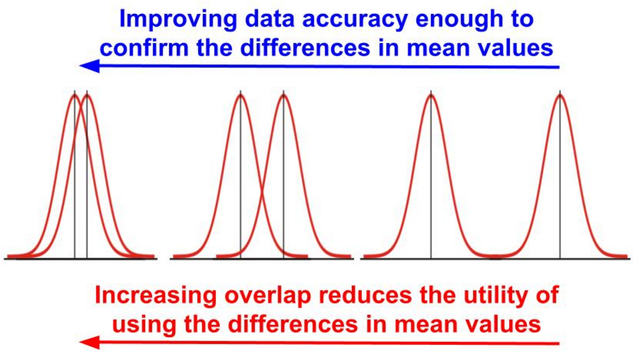
importantly, the probable deviation of an experimental result can be easily known. If the distribution is Gaussian, 68% of measurements will give a result that will have a deviation from the mean of one standard deviation or less. 95% of experimental results will have a deviation of two standard deviations or less.

Heights of male and female adult humans are good examples of Gaussian distributions. {Picture below} The average height of males is 178.4 cm and the standard deviation is 7.6 cm. The average for women is 164.7 cm and the standard deviation is 7.1 cm. Even a brief look is sufficient to know that the distribution for women is less wide, as expected for a distribution with a lower standard deviation. 68% of men are between 170.8 and 186 cm tall. 68% women are between 157.6 and 172.6 cm. Men are taller on an average. The confidence in this statement increases and approaches 100% as we measure the heights of more men and women. If we select one man and one woman at random, is it 100% certain that the man would be taller? No. Because only a few men are taller than all women



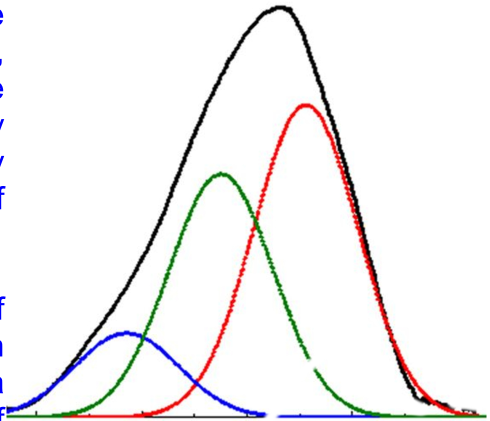
and a few women are shorter than all men. When a man and a woman are selected at random, there will be a finite chance that the woman could be taller. This is called the reverse probability here. The average height of a man is 13.7 cm more than the average height of a woman. The sum of the two standard deviations is 14.7 cm. So the reverse probability will be quite small in this case. Calculating the exact value of the

reverse probability is not possible without advanced mathematics. Stated briefly, as the difference between the two averages or means decreases, the reverse probability increases. If the difference between means is zero, the reverse probability will be 50%.



Consideration of the reverse probability will reduce a lot of friction in any society. Even when the difference between the means of two Gaussian distributions is quite small, by making more measurements and improving the accuracy of measurement, it is possible to confidently assert that the mean of one is larger. But the utility of such a scientific truth decreases. {Picture above} This is extremely important in social issues. For example, imagine that a scientifically valid proof has been presented that on an average one group has higher intelligence coefficient or discipline. This will lead to disputes and discord. But if it is recognized that the reverse probability is very close to 50%, there is no utility of the science and no dispute. If schools and jails are organized on the basis of similar scientific truths, because the reverse probability is close to 50%, the efforts will probably not give the desired results. If the reverse probability is not taken into consideration, there will be a search for people responsible for failure. This will again lead to disputes and discord.

Why are the distributions of the heights of men and women Gaussian? Because the height of an individual depends on a large number of things like the genes, nutrition, history of disease etc. Similarly, the results of many experiments in physics and chemistry are usually

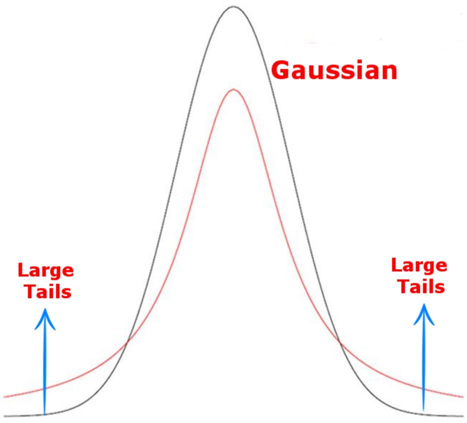


Gaussian because they are dependent on a very large number of things like the ability of the experimenter, environment, age of the instrument, maintenance of the instruments etc. The distribution with each of these may not be Gaussian but it has been proved mathematically that the sum of many distributions is a Gaussian even if each of those distributions is not a Gaussian.

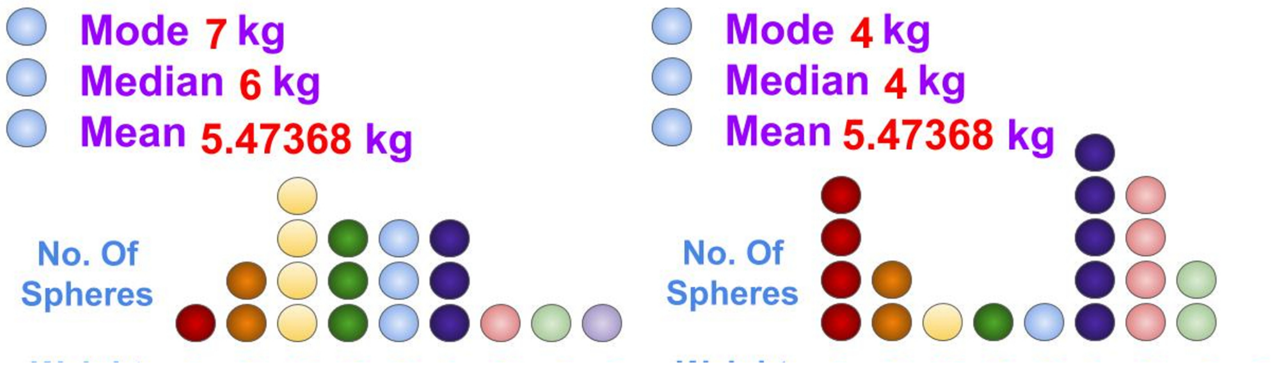
But not all observed distributions are Gaussian. Laws of physics and chemistry are mathematical. So even when an experimentally observed distribution is not really a Gaussian, it can be mathematically divided into a sum of Gaussians. {Picture right} But this is not possible even in medical science or environmental studies though they obey laws of physics and chemistry. There are no mathematical laws relevant to social sciences. It is worth repeating that in physics and chemistry we are dealing with trillions of identical atoms. Thus, it is possible to reduce the standard deviation by improved design of experiments. The standard deviation is most often a limitation of the instruments, not of the atoms. In medicine, environmental studies and social sciences, the standard deviation is quite often not a limitation of the measuring instruments but the complexity of the subject and the difficulty in forming identical groups for study. So in many cases you cannot reduce the standard deviation by altering the experimental procedures.

In a Gaussian distribution, the chance of a measurement with a deviation more than three times the standard deviation is extremely low. For example, in the distribution of heights described before, only one among two thousand males will be taller than 200cm. Many

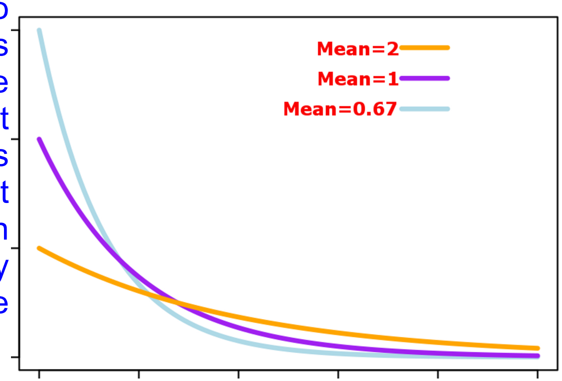
distributions seen in day to day life appear to be approximately Gaussian. But the probabilities for higher multiples of standard deviation are much more than for the Gaussian. These are called flat tail distributions. {Picture right} In particular, these are very common in economic data. Because the deviation for a particularly bad result is many times the standard deviation, experts and political leaders express confidence that there is no danger of such an event. But if the distribution has a fat tail, such events become more likely. The accident really happens and then there is misery all round, followed by searching for the criminals responsible.



If the systems being studied are complex, the distribution can be quite different from a Gaussian. It may even be difficult to compare the distributions. The picture below shows,



two distributions of spheres which were created to illustrate this. The average is the same. There is not much difference in the standard deviation. The median and mode are different. But can we assert that one is superior? Even if the two distributions are known, can they be properly compared? Is it fair to select one because of a superior median value? Any conclusion questionable, particularly since in real life data, the entire distributions are often not known or revealed.



There are some other extremely simple distributions which cannot even be compared theoretically. The exponential distribution is a good example. {Picture above} Calculating the mean and standard deviation in such results is easy but can they be really compared? Such distributions are observed only in some physics and chemistry experiments but not in real life. Taking decisions based on observations in real life is complicated enough even without them.